

# Further implications of the Bessis-Moussa-Villani conjecture

Elliott H. Lieb\* and Robert Seiringer†

\*Departments of Mathematics and Physics, Jadwin Hall,  
Princeton University, P. O. Box 708, Princeton, New Jersey 08544, USA

†Department of Mathematics and Statistics, McGill University,  
805 Sherbrooke St. West, Montreal QC H3A0B9, Canada

June 3, 2012

## Abstract

We find further implications of the BMV conjecture, which states that for hermitian matrices  $\mathbf{A}$  and  $\mathbf{B}$ , the function  $\lambda \mapsto \text{Tr} \exp(\mathbf{A} - \lambda \mathbf{B})$  is the Laplace transform of a positive measure.

Bessis, Moussa and Villani conjectured [1] that for hermitian matrices  $\mathbf{A}$  and  $\mathbf{B}$ ,  $\lambda \mapsto \text{Tr} \exp(\mathbf{A} - \lambda \mathbf{B})$  is the Laplace transform of a positive measure. Previously we showed that this property is equivalent to several other trace inequalities, most notably that the polynomials  $\mathcal{F}_p(\lambda) = \text{Tr} (\mathbf{A} + \lambda \mathbf{B})^p$  for  $p \in \mathbb{N}$  all have positive coefficients when  $\mathbf{A}$  and  $\mathbf{B}$  are positive. Since a proof of the BMV conjecture has recently been put forward by H. Stahl [2], it seems worthwhile to find other implications.

Here we prove two things. Our first result is that the function  $\lambda \in \mathbb{R}_+ \mapsto \text{Tr} (\mathbf{A} + \lambda \mathbf{B})^p$  for general  $p > 0$  has positive derivatives up to order  $[p]$ , the largest integer not less than  $p$ . Moreover, taking further derivatives, one obtains alternating signs for the derivatives. Our second result is that our earlier theorem on  $\mathcal{F}_p(\lambda)$  has a generalization from sums of eigenvalues to elementary symmetric functions of eigenvalues.

---

© 2012 by the authors. This paper may be reproduced, in its entirety, for non-commercial purposes.

# 1 Our previous results

The following theorem was proved in [3].

**THEOREM 1.** *For fixed  $n$  let  $\mathbf{A}$  and  $\mathbf{B}$  denote arbitrary hermitian  $n \times n$  matrices over  $\mathbb{C}$ , and let  $\lambda \in \mathbb{R}$ . The following statements are equivalent:*

- (i) *For all  $\mathbf{A}$  and  $\mathbf{B}$  positive, and all  $p \in \mathbb{N}$ , the polynomial  $\lambda \mapsto \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^p$  has only non-negative coefficients.*
- (ii) *For all  $\mathbf{A}$  hermitian and  $\mathbf{B}$  positive,  $\lambda \mapsto \text{Tr} \exp(\mathbf{A} - \lambda \mathbf{B})$  is the Laplace transform of a positive measure supported in  $[0, \infty)$ .*
- (iii) *For all  $\mathbf{A}$  positive definite and  $\mathbf{B}$  positive, and all  $p \geq 0$ ,  $\lambda \mapsto \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^{-p}$  is the Laplace transform of a positive measure supported in  $[0, \infty)$ .*

We remark that items (i) and (iii) can be combined to the statement that  $\mathcal{F}_p(\lambda) = \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^p$  has positive derivatives when  $p$  is a positive integer, and derivatives of alternating sign when  $p$  is negative.

The positivity property in items (ii) and (iii) follow from Bernstein's Theorem [6] if the functions have alternating derivatives for all  $\lambda \geq 0$ . Since the statement above involves arbitrary  $\mathbf{A}$ , it suffices to check the alternating derivative property at  $\lambda = 0$ .

# 2 Extension to general $p \in \mathbb{R}$

**THEOREM 2.** *Item (ii) in Theorem 1 has the following consequences. For all  $\mathbf{A}$  and  $\mathbf{B}$  positive,  $p \in \mathbb{R}$ , we have*

- a) *For  $1 \leq r \leq [p]$ ,  $\frac{d^r}{d\lambda^r} \mathcal{F}_p(\lambda) \geq 0$  for  $\lambda \geq 0$ .*
- b) *For  $r \geq [p]$  and  $p > 0$ ,  $(-1)^{r-[p]} \frac{d^r}{d\lambda^r} \mathcal{F}_p(\lambda) \geq 0$  for  $\lambda \geq 0$ .*
- c) *For  $r \geq 1$  and  $p \leq 0$ ,  $(-1)^r \frac{d^r}{d\lambda^r} \mathcal{F}_p(\lambda) \geq 0$  for  $\lambda \geq 0$ ,*

where  $\mathcal{F}_p(\lambda) = \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^p$ .

We remark that item c) follows directly from item (iii) in Theorem 1, we included it in Theorem 2 for completeness. Our proof of item b) does not require item (ii), in fact, and we shall give that first.

*Proof of Theorem 2(b).* Let  $s = \lceil p \rceil - p$ . We can assume that  $s > 0$ . We start with the integral representation

$$(\mathbf{A} + \lambda \mathbf{B})^p = \frac{\sin(\pi s)}{\pi} \int_0^\infty \frac{(\mathbf{A} + \lambda \mathbf{B})^{\lceil p \rceil}}{\mathbf{A} + \lambda \mathbf{B} + t} t^{-s} dt. \quad (1)$$

Using the binomial theorem, we have

$$(\mathbf{A} + \lambda \mathbf{B})^{\lceil p \rceil} = \sum_{j=0}^{\lceil p \rceil} \binom{\lceil p \rceil}{j} (-t)^j (\mathbf{A} + \lambda \mathbf{B} + t)^{\lceil p \rceil - j}$$

for  $t > 0$ . In particular, since  $r \geq \lceil p \rceil$ , only the term with  $j = \lceil p \rceil$  contributes to the  $r^{\text{th}}$  derivative of the integrand in (1), i.e.,

$$\frac{d^r}{d\lambda^r} \frac{(\mathbf{A} + \lambda \mathbf{B})^{\lceil p \rceil}}{\mathbf{A} + \lambda \mathbf{B} + t} = \frac{d^r}{d\lambda^r} (-t)^{\lceil p \rceil} \frac{1}{\mathbf{A} + \lambda \mathbf{B} + t}.$$

Hence

$$\frac{d^r}{d\lambda^r} \text{Tr} (\mathbf{A} + \lambda \mathbf{B})^p = (-1)^{\lceil p \rceil} \frac{\sin(\pi s)}{\pi} \int_0^\infty \text{Tr} \frac{d^r}{d\lambda^r} \frac{1}{\mathbf{A} + \lambda \mathbf{B} + t} t^p dt.$$

Using the resolvent identity, we have

$$\frac{d^r}{d\lambda^r} \frac{1}{\mathbf{A} + \lambda \mathbf{B} + t} = \frac{(-1)^r}{\mathbf{A} + \lambda \mathbf{B} + t} \left( \mathbf{B} \frac{1}{\mathbf{A} + \lambda \mathbf{B} + t} \right)^r,$$

from which we easily conclude that

$$(-1)^r \text{Tr} \frac{d^r}{d\lambda^r} \frac{1}{\mathbf{A} + \lambda \mathbf{B} + t} \geq 0,$$

which completes the proof.  $\square$

For the proof of part a) of Theorem 2, we shall need the following lemma.

**LEMMA 1.** *Let  $\mathbf{a}$  and  $\mathbf{b}$  be hermitian  $n \times n$  matrices over  $\mathbb{C}$ , with  $\mathbf{a}$  positive definite. Define  $\mathbf{A} = \mathbf{a}^{-1}$  and  $\mathbf{B} = \mathbf{a}^{-1/2} \mathbf{b} \mathbf{a}^{-1/2}$ , and let  $\lambda \in \mathbb{R}$ . For all  $p \in \mathbb{C}$  and  $r \in \mathbb{N}$*

$$(p+r) \frac{d^r}{d\lambda^r} \text{Tr} \frac{1}{(\mathbf{a} + \lambda \mathbf{b})^p} \Big|_{\lambda=0} = p(-1)^r \frac{d^r}{d\lambda^r} \text{Tr} (\mathbf{A} + \lambda \mathbf{B})^{p+r} \Big|_{\lambda=0}. \quad (2)$$

*Proof.* The proof of (2) for  $p \in \mathbb{N}$  was given in [3]; we include it verbatim in the appendix for completeness.

Both sides of (2) are entire functions of  $p$ . Let  $f(p)$  denote the left side minus the right side. We have just noted that  $f(p) = 0$  for  $p \in \mathbb{N}$ . Moreover, we can find an  $a > 0$  such that the function  $p \mapsto f(p)e^{-ap}$  is bounded for  $\Re p \geq 0$ . It then follows from Carlson's theorem (see [4]) that  $f$  is identically zero in the half-space  $\Re p \geq 0$ , and hence for all  $p \in \mathbb{C}$ .  $\square$

*Proof of Theorem 2(b).* As remarked after Theorem 1, it is sufficient to prove the statement for  $\lambda = 0$ . We can assume that  $p > r$ , the statement is trivial for  $p = r$ . From the identity (2) with  $p$  replaced by  $p - r$ , we have

$$\left. \frac{d^r}{d\lambda^r} \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^p \right|_{\lambda=0} = (-1)^r \frac{p}{p-r} \left. \frac{d^r}{d\lambda^r} \text{Tr}(\mathbf{a} + \lambda \mathbf{b})^{r-p} \right|_{\lambda=0}. \quad (3)$$

By item (iii) of Theorem 1, the function  $\lambda \mapsto \text{Tr}(\mathbf{a} + \lambda \mathbf{b})^{r-p}$  has alternating derivatives for  $p > r$ , and hence the right side of (3) is positive.  $\square$

### 3 Extension to elementary symmetric functions of eigenvalues

We now return to the polynomials  $\mathcal{F}_p(\lambda) = \text{Tr}(\mathbf{A} + \lambda \mathbf{B})^p$  for *positive integer*  $p$ . The coefficient of  $\lambda^k$  in this polynomial is the trace of a sum of  $p$ -letter words in two letters, with  $\mathbf{A}$  appearing  $p - k$  times, and  $\mathbf{B}$  appearing  $k$  times. It is known that the trace of an individual words need not be positive [7], but the sum of the traces is, according to Theorem 1 and the BMV conjecture.

Instead of traces, which involve sums of eigenvalues, let us consider determinants, which are the products of all the eigenvalues. It is clear that the determinant of the sum of all words for given  $p$  and  $k$  need not be positive, as the example  $p = 2$  and  $k = 1$  shows; namely, while  $\text{Tr}(\mathbf{AB} + \mathbf{BA})$  is positive, the determinant  $\det(\mathbf{AB} + \mathbf{BA})$  need not be [5]. On the other hand, each individual determinant is clearly positive, since it is equal to  $(\det \mathbf{A})^{p-k}(\det \mathbf{B})^k$ . This suggests that general elementary symmetric functions of the eigenvalues of the sum of all words need not be positive, while the sum of the elementary symmetric functions of the eigenvalues of the individual words should be positive. We prove that here.

The elementary symmetric functions of the eigenvalues, of which the trace and the determinant are two special cases, are given by

$$e_j(\lambda_1, \dots, \lambda_n) = \sum_{1 \leq i_1 < i_2 < \dots < i_j \leq n} \prod_{k=1}^j \lambda_{i_k}$$

for  $1 \leq j \leq k$ . This number  $e_j$  is also equal to the sum of the principal subdeterminants of order  $j$  of a matrix  $\mathbf{M}$  with eigenvalues  $\lambda_1, \dots, \lambda_n$ , and we shall also write  $e_j(\mathbf{M})$  for short. Another way to think of  $e_j(\mathbf{M})$  is as the trace of the  $j^{\text{th}}$  adjugate of  $\mathbf{M}$ , which is the  $j$ -fold anti-symmetric tensor product of  $\mathbf{M}$  with itself,  $\mathbf{M} \wedge \dots \wedge \mathbf{M}$ .

**THEOREM 3.** *Item (i) in Theorem 1, assumed to hold for all  $n \in \mathbb{N}$ , has the following consequence. For all  $\mathbf{A}$  and  $\mathbf{B}$  positive,  $1 \leq k \leq p$  and  $1 \leq j \leq n$ ,*

$$\sum_{i=1}^{\binom{p}{k}} e_j(\mathbf{W}_i) \geq 0$$

where the  $\mathbf{W}_i$  denote all words of length  $p$  with  $k$  letters  $\mathbf{B}$  and  $p - k$  letters  $\mathbf{A}$ .

As remarked above, in general it is false that  $e_j(\sum_i \mathbf{W}_i) \geq 0$ , except for  $j = 1$ , where  $\sum_i e_j(\mathbf{W}_i) = e_j(\sum_i \mathbf{W}_i)$ .

*Proof.* We apply item (i) in Theorem 1 to the  $j$ -fold antisymmetric tensor products of  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. Since the tensor product of a product of matrices equals the product of the individual tensor products, the theorem follows immediately.  $\square$

In a similar fashion, we can prove the following.

**THEOREM 4.** *Item (ii) in Theorem 1, assumed to hold for all  $n \in \mathbb{N}$ , has the following consequences.*

a) *For all hermitian  $\mathbf{A}$  and positive  $\mathbf{B}$ ,  $k \geq 1$  and  $1 \leq j \leq n$ ,*

$$\int_{s_i \geq 0, \sum_{i=1}^{k+1} s_i = 1} e_j(e^{s_1 \mathbf{A}} \mathbf{B} e^{s_2 \mathbf{A}} \mathbf{B} \dots \mathbf{B} e^{s_{k+1} \mathbf{A}}) ds_1 \dots ds_{k+1} \geq 0$$

b) For all hermitian  $A$  and positive  $B$ , and  $1 \leq j \leq n$ ,

$$\lambda \mapsto e_j(e^{A-\lambda B})$$

is the Laplace transform of a positive measure supported in  $[0, \infty)$ .

*Proof.* Item a) follows by applying item (ii) with  $A$  replaced by

$$\alpha = (A \wedge \mathbb{1} \wedge \cdots \wedge \mathbb{1}) + (\mathbb{1} \wedge A \wedge \mathbb{1} \cdots \wedge \mathbb{1}) + \cdots + (\mathbb{1} \wedge \cdots \wedge \mathbb{1} \wedge A)$$

and  $B$  replaced by  $\beta = B \wedge B \wedge \cdots \wedge B$ . Note that  $e^\alpha = e^A \wedge e^A \wedge \cdots \wedge e^A$ . The  $k^{\text{th}}$  derivative of  $e^{\alpha-\lambda\beta}$  with respect to  $\lambda$  is equal to

$$(-1)^k \int_{s_i \geq 0, \sum_{i=1}^{k+1} s_i = 1} e^{s_1 \alpha} \beta e^{s_2 \alpha} \beta \cdots \beta e^{s_{k+1} \alpha} ds_1 \cdots ds_{k+1}$$

and hence the statements follows.

To obtain item b), we replace  $A$  by  $\alpha$  and  $B$  by

$$\gamma = (B \wedge \mathbb{1} \wedge \cdots \wedge \mathbb{1}) + (\mathbb{1} \wedge B \wedge \mathbb{1} \cdots \wedge \mathbb{1}) + \cdots + (\mathbb{1} \wedge \cdots \wedge \mathbb{1} \wedge B).$$

Then  $e^{\alpha-\lambda\gamma} = e^{A-\lambda B} \wedge e^{A-\lambda B} \wedge \cdots \wedge e^{A-\lambda B}$ .  $\square$

The case  $j = n$  in item b) of Theorem 4 follows immediately from the fact that

$$\det e^{A-\lambda B} = e^{\text{Tr } A - \lambda \text{Tr } B},$$

and holds even without the assumption of the Theorem.

Note that

$$e_2(M) = \frac{1}{2} ((\text{Tr } M)^2 - \text{Tr } M^2).$$

For  $M = e^{A-\lambda B}$ , both  $(\text{Tr } M)^2$  and  $\text{Tr } M^2$  are the Laplace transform of a positive measure. It is remarkable that also their difference has this property! Similar conclusions can be drawn for general  $j \geq 2$ .

## A Appendix: Proof of Lemma 1 for $p \in \mathbb{N}$

By induction it is easy to show that

$$\frac{d^r}{d\lambda^r} (A + \lambda B)^{p+r} = r! \sum_{\substack{0 \leq i_1, \dots, i_{r+1} \leq p \\ \sum_j i_j = p}} (A + \lambda B)^{i_1} B \cdots B (A + \lambda B)^{i_{r+1}}. \quad (4)$$

By taking the trace at  $\lambda = 0$  we obtain

$$I_1 \equiv \left. \frac{d^r}{d\lambda^r} \text{Tr} (A + \lambda B)^{p+r} \right|_{\lambda=0} = r! \sum_{\substack{0 \leq i_1, \dots, i_{r+1} \leq p \\ \sum_j i_j = p}} \text{Tr} A^{i_1} B \dots B A^{i_{r+1}} . \quad (5)$$

Moreover, by similar arguments,

$$\frac{d^r}{d\lambda^r} \frac{1}{(a + \lambda b)^p} = (-1)^r r! \sum_{\substack{1 \leq i_1, \dots, i_{r+1} \leq p \\ \sum_j i_j = p+r}} \frac{1}{(a + \lambda b)^{i_1}} b \dots b \frac{1}{(a + \lambda b)^{i_{r+1}}} . \quad (6)$$

By taking the trace at  $\lambda = 0$  and using cyclicity, we get

$$I_2 \equiv \left. \frac{d^r}{d\lambda^r} \text{Tr} \frac{1}{(a + \lambda b)^p} \right|_{\lambda=0} = (-1)^r r! \sum_{\substack{0 \leq i_1, \dots, i_{r+1} \leq p-1 \\ \sum_j i_j = p-1}} \text{Tr} A A^{i_1} B \dots B A^{i_{r+1}} . \quad (7)$$

We have to show that

$$I_2 = \frac{p}{p+r} (-1)^r I_1 . \quad (8)$$

To see this we rewrite  $I_1$  in the following way. Define  $p+r$  matrices  $M_j$  by

$$M_j = \begin{cases} B & \text{for } 1 \leq j \leq r \\ A & \text{for } r+1 \leq j \leq r+p . \end{cases} \quad (9)$$

Let  $\mathcal{S}_n$  denote the permutation group. Then

$$I_1 = \frac{1}{p!} \sum_{\pi \in \mathcal{S}_{p+r}} \text{Tr} \prod_{j=1}^{p+r} M_{\pi(j)} . \quad (10)$$

Because of the cyclicity of the trace we can always arrange the product such that  $M_{p+r}$  has the first position in the trace. Since there are  $p+r$  possible locations for  $M_{p+r}$  to appear in the product above, and all products are equally weighted, we get

$$I_1 = \frac{p+r}{p!} \sum_{\pi \in \mathcal{S}_{p+r-1}} \text{Tr} A \prod_{j=1}^{p+r-1} M_{\pi(j)} . \quad (11)$$

On the other hand,

$$I_2 = (-1)^r \frac{1}{(p-1)!} \sum_{\pi \in \mathcal{S}_{p+r-1}} \text{Tr } \mathbf{A} \prod_{j=1}^{p+r-1} M_{\pi(j)} , \quad (12)$$

so we arrive at the desired equality.

*Acknowledgments.* We are grateful to P. Landweber for valuable discussion on this topic, especially about elementary symmetric functions. Partial financial support by U.S. NSF grant PHY-0965859 (E.H.L.) and the NSERC (R.S.) is gratefully acknowledged.

## References

- [1] D. Bessis, P. Moussa and M. Villani, *Monotonic converging variational approximations to the functional integrals in quantum statistical mechanics*, J. Math. Phys. **16**, 2318–2325 (1975).
- [2] H.R. Stahl, *Proof of the BMV Conjecture*, preprint, arXiv:1107.4875
- [3] E.H. Lieb, R. Seiringer, *Equivalent forms of the Bessis-Moussa-Villani conjecture*, J. Stat. Phys. **115**, 185–190 (2004).
- [4] G.H. Hardy, *On two theorems of F. Carlson and S. Wigert*, Acta Math. **42**, 327–339 (1920).
- [5] C.S. Ballantine, *A Note on the Matrix Equation  $H = AP + PA^*$* , Linear Algebra Appl. **2**, 37–47 (1969).
- [6] W. F. Donoghue, *Monotone matrix functions and analytic continuation*, Springer (1974).
- [7] C. R. Johnson and C. J. Hillar, *Eigenvalues of words in two positive definite letters*, SIAM J. Matrix Anal. Appl. **23**, 916–928 (2002).